# A Novel Technique of Rainfall bias Correction using Monthly Hybrid Approach

**Deepak Singh Bisht[1,2*], Chandranath Chatterjee[2], Narendra Singh Raghuwanshi[2,3]**

*[1]National Institute of Hydrology, Western Himalayan Regional Centre, Jammu*

*[2]Department of Agricultural and Food Engineering, Indian Institute of Technology Kharagpur, Kharagpur, India*

*[3]Maulana Azad National Institute of Technology, Bhopal, India*

*Corresponding author email Id: dsbisht.ae@gmail.com*

**Abstract:** Rainfall projections from Global Climate Models (GCMs) aid in assessing the water resources availability and dynamics in unforeseen future. However, these projections are often marred by biases, possibly due to imprecise model assumptions or model's inability in capturing geophysical processes with greater accuracy. Therefore, it is imperative to treat the rainfall projections obtained from GCMs using suitable techniques to remove the biases prior to their use in subsequent studies. Present study proposes a novel 'monthly hybrid approach' for rainfall bias correction and compares it with a traditional bias correction technique i.e., 'seasonal approach' and a comparatively newer 'hybrid approach'. Unlike monthly hybrid approach which explicitly operates at monthly scale, hybrid approach operates at annual scale whereas seasonal approach utilizes three sub-sample time series covering JJAS, ONDJF, and MAM months. Rainfall projections from a total of nine GCMs from CMIP5 family of models were utilized to perform the comparison using quantile mapping approach. Gamma and Generalized Extreme Value (GEV) distributions were used in hybrid and monthly hybrid approaches (details are discussed in the paper) for bias correction of normal and extreme events, respectively over Mahanadi river basin. Quarter degree gridded rainfall time series for the time period of 1951-2005 obtained from India Meteorological Department (IMD) was used as reference dataset for bias-correction. Prior to bias-correction, frequency correction for rainy days in GCM projections was carried out for the training period covering the time span of 30 years during 1976-2005. For frequency correction of rain days, IMD rainfall time series during training period was used as benchmark. Threshold values of rainfall during training period for each grid was estimated and applied for respective GCMs for correcting the rainy days. These threshold values were further used for testing period i.e., 1951-1975 covering 25 years of time span to correct the rainy days frequency. Suitability of bias correction techniques for different GCMs were assessed by comparing regional rainfall cycle obtained from bias corrected time series with reference rainfall cycle produced using IMD data for training and testing period. Furthermore, the overall accuracy achieved by different techniques were assessed using Taylor's diagram that includes the root mean square difference, standard deviation, and correlation coefficient. The efficacy of bias-correction techniques in resolving extreme rainfall were evaluated using l-moments based frequency analysis of bias-corrected and reference (IMD) rainfall time series. Proposed bias-correction technique using monthly hybrid approach is found to be more efficient in resolving rainfall climatology over the region as well as the extreme events. Therefore, it is recommended to perform bias correction approach using monthly hybrid approach to achieve higher accuracy.

**Keywords:** Bias-correction; Rainfall; Quantile mapping; Extreme events; L-moments

## 1 Introduction

As per the IPCC (2007), the frequency of extreme rainfall events has increased in last 4-5 decades. In such case, it becomes imperative to analyze potential flood risks to minimize the economic losses by adapting suitable preventive measures. Rainfall projections from GCMs can be utilized to analyze the expected extremities in future time period in a hydrological modeling framework. Rainfall, being the key input in simulating the runoff response of a basin is very critical, however, the rainfall projections from GCMs are available at coarser scale and cannot be directly applied to regional studies. This necessitates the need to downscale the coarser GCM rainfall projections at local level for improved applicability. Statistical downscaling techniques are applied for this purpose in many studies (Wilby and Wigley, 1997; Um et al., 2016; Dhage et al, 2016; Salvi et al., 2013; Teutschbein and Seibert, 2012). Quantile mapping is one of the most widely used technique for bias correction of GCM projections. It has been reported that compared to annual scale, quantile mapping at sub-annual time scales i.e., semi-annual, seasonal or monthly time scale results into improved performance (Reiter et al., 2017). In a recent study (Um et al., 2016), a new approach of bias-correction, named Hybrid approach, that corrects extreme precipitation explicitly found to yield better results. To correct the rainfall extremes along with the rainfall seasonal cycle is critical to utilize the GCM projection appropriately in hydrological modeling or climate change impact analysis. Therefore, in present study an attempt has been made to identify the most appropriate method of rainfall bias-correction from three selected approaches (1) seasonal approach (2) monthly approach, and (3) monthly hybrid approach. Here, it is worth noting that though seasonal and monthly approaches have often been reported to resolve the rainfall climatology, (Reiter et al., 2017), their skills in correcting extreme events have not been subjected to evaluation. Quantile mapping using specific distribution function for normal rainfall and extreme rainfall i.e., hybrid approach as proposed by Um et al. (2016), is reported to resolve the extreme with improved efficiency. However, Um et al. (2016) presented the study for annual time series, which may not capture the rainfall seasonality in monsoon dominated regions of India. To overcome this, a comparatively new approach, i.e., monthly hybrid approach that applies month-wise quantile mapping on normal rainfall using gamma distribution with explicit correction of extreme rainfall using GEV distribution is proposed. A comparison of afore-discussed three approaches in capturing the rainfall extremes and seasonal cycle has been presented in the paper and discussed in detail in subsequent sections. The study utilizes a total of 9 GCMs to perform the comparison and utilizes l-moments (Hosking, 1990) based frequency analysis approach to estimate 50-year return period rainfall for comparison of corrected extremes. In present study, the entire analysis is carried out for Mahanadi river basin, which can be replicated for any other river basin in similar fashion.

The paper is organized as follows. Section 2 describes the material and methods. Results and discussion are presented in sections 3, and section 4 presents the conclusions derived from the study.

## 2 Materials and Methods

### 2.1 Data used and study area

Gridded rainfall (Pai et al., 2014) at quarter degree spatial resolution, procured from IMD, was utilized in the present study to perform the quantile mapping on the simulated precipitation of nine GCMs, namely BCC-CSM1.1(m), HadGEM2-AO, GFDL-CM3, GFDL-ESM2G, IPSL-CM5A-LR, IPSL-CM5A-MR, MIROC5, MIROC-ESM-CHEM, NorESM1-M. GCM data was downloaded from Earth System Grid Federation web portal (https://esgf-node.llnl.gov/projects/esgf-llnl/). Following Bisht et al., (2019) and Akhter et al., (2017) GCM rainfall was remapped at quarter degree spatial resolution using Climate Data Operator package to make it spatially consistent with IMD grids.

The study was performed over Mahanadi River basin covering a geographically area of 1,41,589 $km^2$. A total of 264 grids of quarter degree resolution over Mahanadi were considered for comparison of different bias-correction approaches. Average annual rainfall of the basin is around 1572 mm which is dominated by south-west monsoon season.

### 2.2 Bias-correction of GCM rainfall

In present study, simulated GCM daily rainfall was bias corrected employing quantile mapping method (Li et al., 2010) as illustrated in Dhage et al. (2016) and Salvi et al.(2013). To perform bias-correction of daily rainfall, historical time series of GCM data was divided into training (1976-2005 i.e., 30 years for calibration) and testing (1951-1975, i.e., 25 years for validation) periods. In seasonal approach, quantile mapping was performed on daily time series of seasonal rainfall that comprises monsoon (JJAS), post-monsoon (ONDJF), and pre-monsoon (MAM) months using gamma distribution following (Li et al., 2010; Reiter et al., 2017; Salvi et al., 2013; Teutschbein and Seibert, 2012; Um et al., 2016). In another method, i.e., hybrid approach as proposed by (Um et al., 2016), annual maximum rainfall was bias-corrested using GEV distribtuion whereas the normal rainfall time series i.e., after removing annual maximum rainfall, was corrected using gamma distribtuion. However, while performing the bias-correction it is very likely that quantile mapping involving gamma distribution may yield higher rainfall values in comparison to the corrected extreme values for annual maximum series using GEV distribution, also observed in (Um et al., 2016). To overcome this, values having higher magnitude in the series corrected using gamma distribution in comparison to bias-corrected annual maximum values for a particular year were treated as missing values and computed using interpolation as presented in Um et al. (2016). In present study, Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) interpolation technique was used to obtain the interpolated values owing to the fact that it has no overshooting problems and less oscillations if the data is not smooth. Owing to the fact, that hybrid approach operates at annual scale it is likely to be less efficient in capturing the seasonal pattern of rainfall of monsoon dominated climate of Indian region. Seasonal climatology can be efficiently captured while performing quantile mapping on a monthly or sub-annual time series (Reiter et al., 2017). Therefore, a modification in hybrid approach is proposed in rainfall bias correction wherein annual maximum values were corrected using GEV distribution and quantile mapping using gamma distribution was performed on month-wise daily time series instead of year-wise daily time series. Mathematical expressions for quantile mapping techniques are discussed elsewhere (Dhage et al., 2016; Salvi et al., 2013; Um et al., 2016) therefore, not included in the present study.

### 2.3 Frequency analysis of extreme rainfall

L-moments (Hosking, 1990) technique to identified the best fit distribution employs multiple distributions to obtain the parameters for best fitted distribution to the dataset. It uses, five 3-parameter distributions i.e., generalized extreme value (GEV), generalized logistic (GLO), generalized normal (GNO), Pearson type-III (PE3), generalized pareto (GPA), and one 5-parameter distribution Wakeby (WAK). Estimated best fitted parameters thereafter, used for computing the frequency factor to estimate the values of different return periods. In present study, L-moments based (Hosking, 1990) frequency analysis method was used to derive the 50-year return period using annual maximum rainfall values for observed and bias corrected rainfalls. L-moments technique has been reported to perform better over other methods of parameter estimation such as, method of moments, maximum likelihood, probability weighted moments (Hosking and Wallis, 1997; Kumar and Chatterjee, 2011); and has been widely used in similar studies for frequency analysis (Jacob et al., 2019; Jena et al., 2014; Kumar et al., 2003; Kumar and Chatterjee, 2005; Samantaray et al., 2015). Readers can refer Kumar and Chatterjee (2011) for detailed description on l-moments method.

## 3 Results and Discussion

Rainfall data of all the nine GCMs for retrospective period i.e., 1951-2005 was bias corrected using afore-discussed three approaches, namely seasonal, hybrid, and monthly hybrid. Prior to the application of quantile mapping, frequency correction for rainfall days were performed on month-wise basis, that essentially tries to ensure realistic distribution of rainy days across the months in GCM simulated rainfall time series. To apply this correction, grid-wise threshold values were obtained for each month during training period i.e., 1976-2005. To estimate the thresholds, IMD and GCM time series of respective grids were arranged in descending order and the corresponding value of GCM for the sequence below which IMD time series shows no rainfall was identified. This threshold then applied for testing period i.e., 1951-1975 for frequency correction. The bias-correction approaches were evaluated against each other for their skills by comparing regional rainfall pattern resolved by bias-corrected rainfalls with IMD rainfall (Fig. 1). Besides, suitability of bias-corrected techniques were also assessed by comparing the pattern of bias-corrected rainfall with IMD data in terms of variation, i.e.

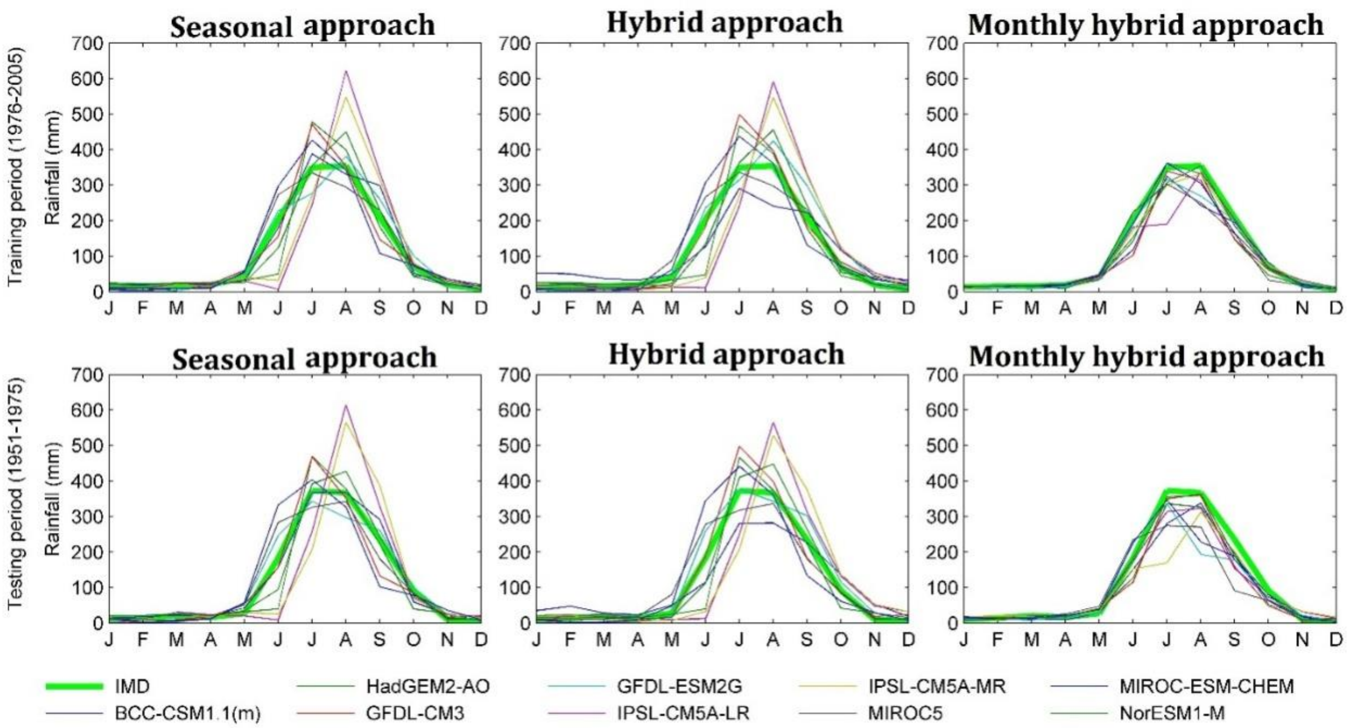


**Fig. 1** Comparison of mean monthly rainfall cycle of bias-corrected rainfall of 9 GCMs using seasonal, hybrid, and monthly hybrid approaches with IMD data(x-axis is for months)

standard deviation (SD), correlation coefficient (CC), and root-mean-square difference (RMSD) utilizing Taylor diagram (Taylor, 2001) shown in Fig. 2. As the monthly hybrid approach applies the quantile mapping on month-wise daily time series unlike seasonal and hybrid approaches that perform bias-correction on daily time series of seasonal and annual timescales, respectively, monthly hybrid approach captures the rainfall cycle with improved accuracy as shown in Fig. 1 for majority of the GCMs studied. Though it was observed that monthly hybrid approach shows underestimation in rainfall compared to other two approaches (Table 1), it produces smaller Root mean Square Error (RMSE) and Mean Absolute Error (MAE) for majority of GCMs. Similarly, it also shows lesser departure from observation (Fig. 2) in comparison to seasonal and hybrid approaches during training and testing period. Thus, monthly hybrid approach provides improvement in capturing regional rainfall cycle. Besides, resolving seasonal pattern over regional scale bias-correction approaches were further evaluated for their skills in correcting extreme rainfall events by comparing 50-year return period rainfall estimated using l-moments based frequency analysis approach as discussed earlier. Here, it is important to reiterate that hybrid approach as well as monthly hybrid approach correct the annual maximum rainfall explicitly using GEV distribution unlike seasonal rainfall, therefore, both yield similar results for extremes. Hence, the 50-year return period rainfall using l-moments based frequency analysis was estimated only for seasonal and monthly hybrid approaches for comparison.

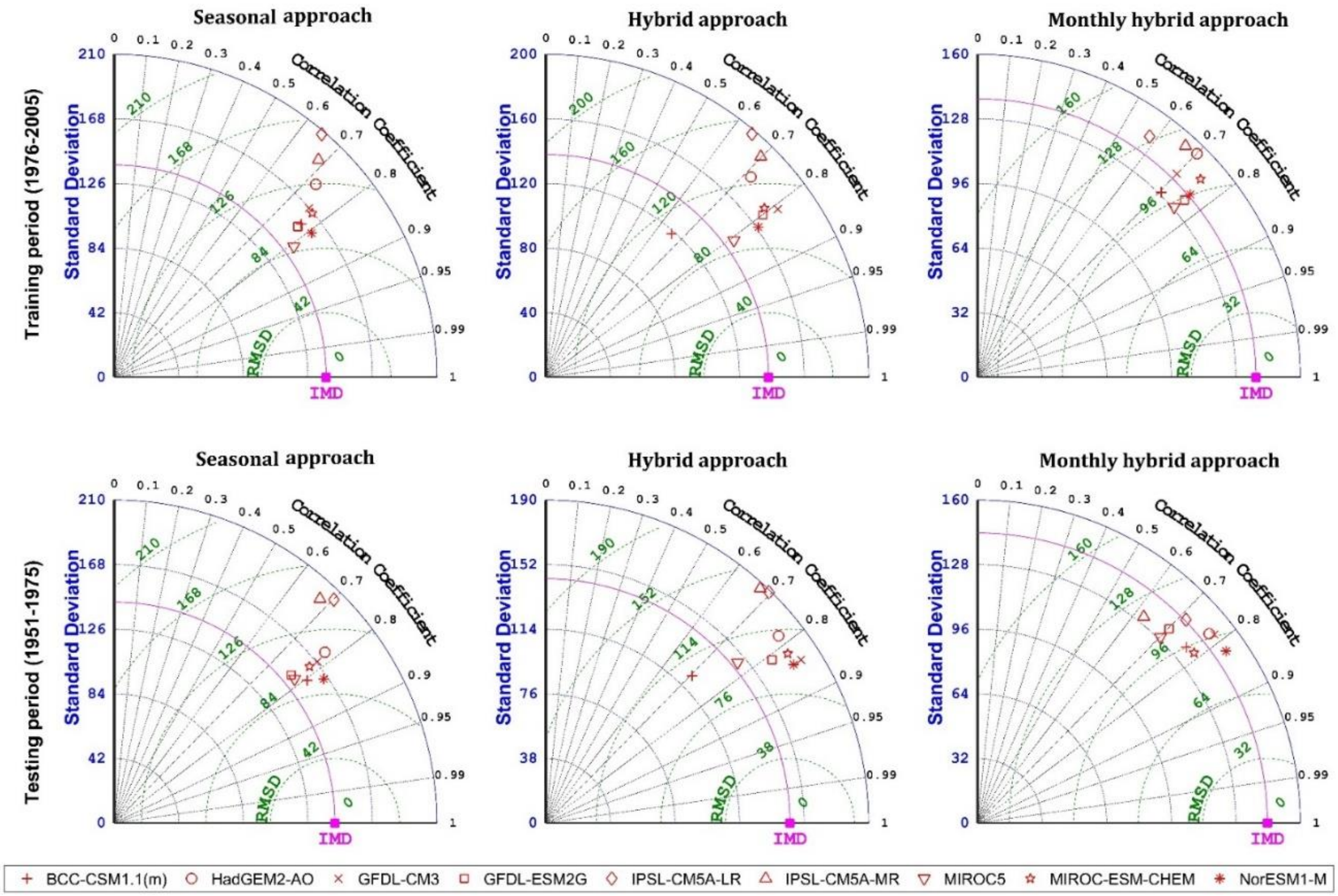


**Fig. 2** Taylor diagram for statistical comparison of year wise regional monthly IMD (Observed) precipitation with bias-corrected rainfall of 9 GCMs using seasonal, hybrid, and monthly hybrid approaches

**Table 1** Comparison of performance evaluation statistics of seasonal, hybrid, and monthly hybrid approaches in producing regional rainfall cycle

| Period | Metric | GCM / Approach | BCC-CSM1.1 (m) | HadGEM2-AO | GFDL-CM3 | GFDL-ESM2G | IPSL-CM5A-LR | IPSL-CM5A-MR | MIROC5 | MIROC-ESM-CHEM | NorESM1-M |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Training | PBias | Seasonal | 6.28 | -0.19 | 1.90 | 2.51 | 7.99 | 4.58 | 1.42 | 5.67 | 4.09 |
| | | Hybrid | -1.42 | -1.46 | 9.55 | 11.02 | 9.62 | 7.90 | 3.34 | 10.30 | 7.36 |
| | | Monthly Hybrid | -17.95 | -11.96 | -16.42 | -12.11 | -14.99 | -9.07 | -14.95 | -6.59 | -9.34 |
| | RMSE (mm) | Seasonal | 30.82 | 60.40 | 42.77 | 30.35 | 107.38 | 84.50 | 28.02 | 46.37 | 38.29 |
| | | Hybrid | 49.72 | 58.57 | 46.87 | 40.94 | 101.98 | 86.44 | 29.31 | 46.32 | 38.83 |
| | | Monthly Hybrid | 41.70 | 27.13 | 36.25 | 26.78 | 47.00 | 19.69 | 37.44 | 19.67 | 20.50 |
| | MAE (mm) | Seasonal | 18.94 | 35.45 | 26.75 | 21.84 | 62.43 | 50.64 | 17.69 | 30.84 | 24.20 |
| | | Hybrid | 40.16 | 34.31 | 28.58 | 30.92 | 67.66 | 59.24 | 19.02 | 31.03 | 26.07 |
| | | Monthly Hybrid | 21.98 | 15.46 | 19.73 | 13.25 | 19.35 | 12.39 | 22.31 | 10.81 | 11.50 |
| Testing | PBias | Seasonal | 2.71 | -3.00 | -2.27 | -0.24 | 4.57 | 2.35 | -1.54 | 1.92 | 1.34 |
| | | Hybrid | -4.13 | -4.60 | 6.02 | 7.64 | 6.12 | 5.48 | 0.04 | 7.38 | 4.45 |
| | | Monthly Hybrid | -17.28 | -14.45 | -13.06 | -18.33 | -11.77 | -21.56 | -20.68 | -13.15 | -6.22 |
| | RMSE (mm) | Seasonal | 21.15 | 52.83 | 42.60 | 31.46 | 97.74 | 98.30 | 37.24 | 60.97 | 31.90 |
| | | Hybrid | 47.36 | 52.15 | 41.40 | 32.62 | 88.63 | 91.88 | 39.65 | 60.02 | 34.47 |
| | | Monthly Hybrid | 34.34 | 28.04 | 32.90 | 56.02 | 33.09 | 63.16 | 60.29 | 45.54 | 16.22 |
| | MAE (mm) | Seasonal | 13.67 | 31.29 | 26.95 | 23.13 | 59.39 | 61.81 | 24.16 | 38.52 | 17.85 |
| | | Hybrid | 37.74 | 30.16 | 25.01 | 23.13 | 58.97 | 65.71 | 26.44 | 37.36 | 21.74 |
| | | Monthly Hybrid | 21.19 | 20.12 | 20.72 | 31.02 | 22.02 | 34.40 | 38.85 | 27.16 | 9.76 |

Best-fit distribution for each grids were estimated using l-moments approach for annual maximum rainfall time series of IMD, bias-corrected GCMs using both the approaches for training (1976-2005) an testing periods (1951-1975). These grid-wise best fit parameters were then utilized in computing frequency factor for 50-year return period rainfall. Mathematical expression and formulations can be referred in Kumar and Chatterjee (2011) for improved understanding which are not included in this paper keeping the space limitation in view.

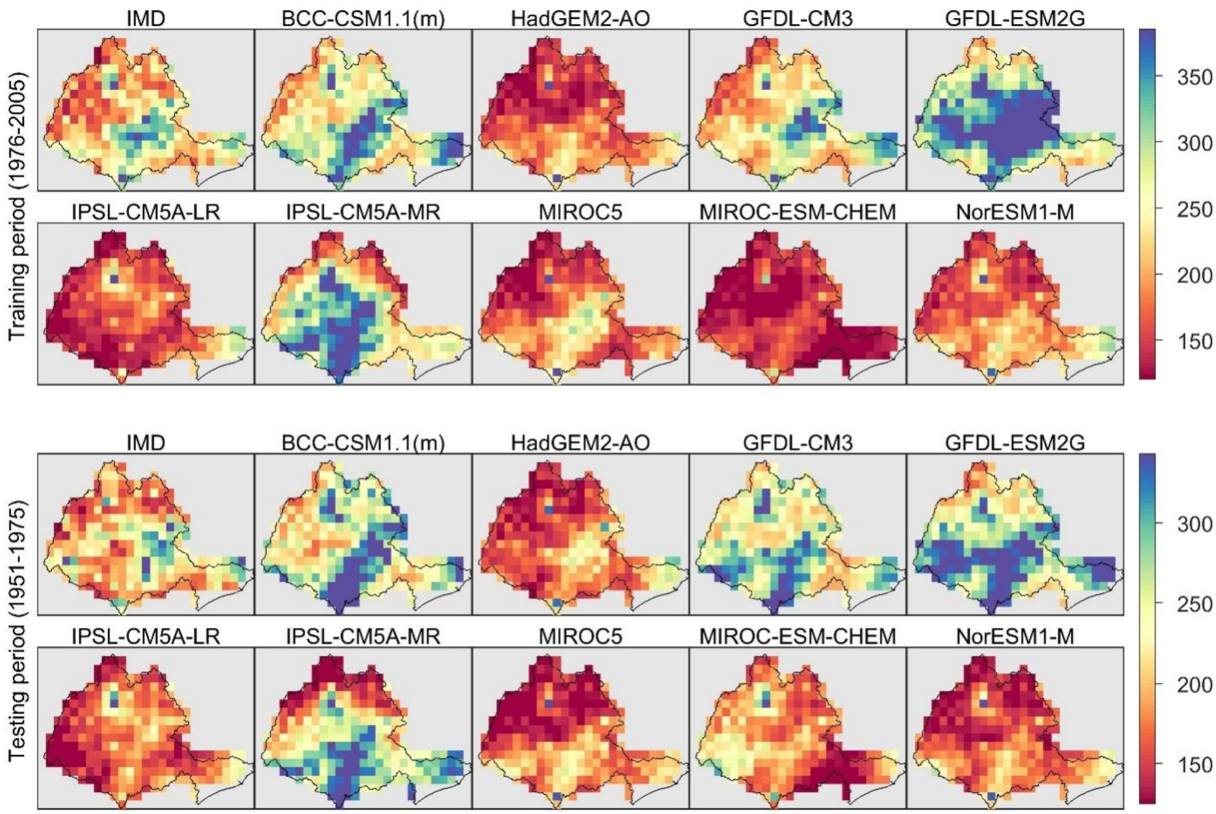


**Fig. 3** Spatial maps of 50 year return period rainfall for IMD and bias-corrected GCMs using seasonal approach. Color bar in the right shows rainfall depth in mm.

As evident from Fig. 3, bias-corrected rainfall values of 50-year return period using seasonal approach do not correlates well spatially with the 50-year return period rainfall estimates obtained using IMD data. For both training and testing period, seasonal bias correction technique which does not corrects the extremes explicitly found to be perform poorly. On the contrary, the monthly hybrid approach as illustrated in Fig. 4 shows close agreement with the IMD estimates across the grids for both training (1976-2005) and testing periods (1951-2005). It is worth to note that, for majority of the GCMs monthly hybrid approach also showed improvement in RMSE and MAE during training and testing periods in comparison to seasonal approach (Table 1).

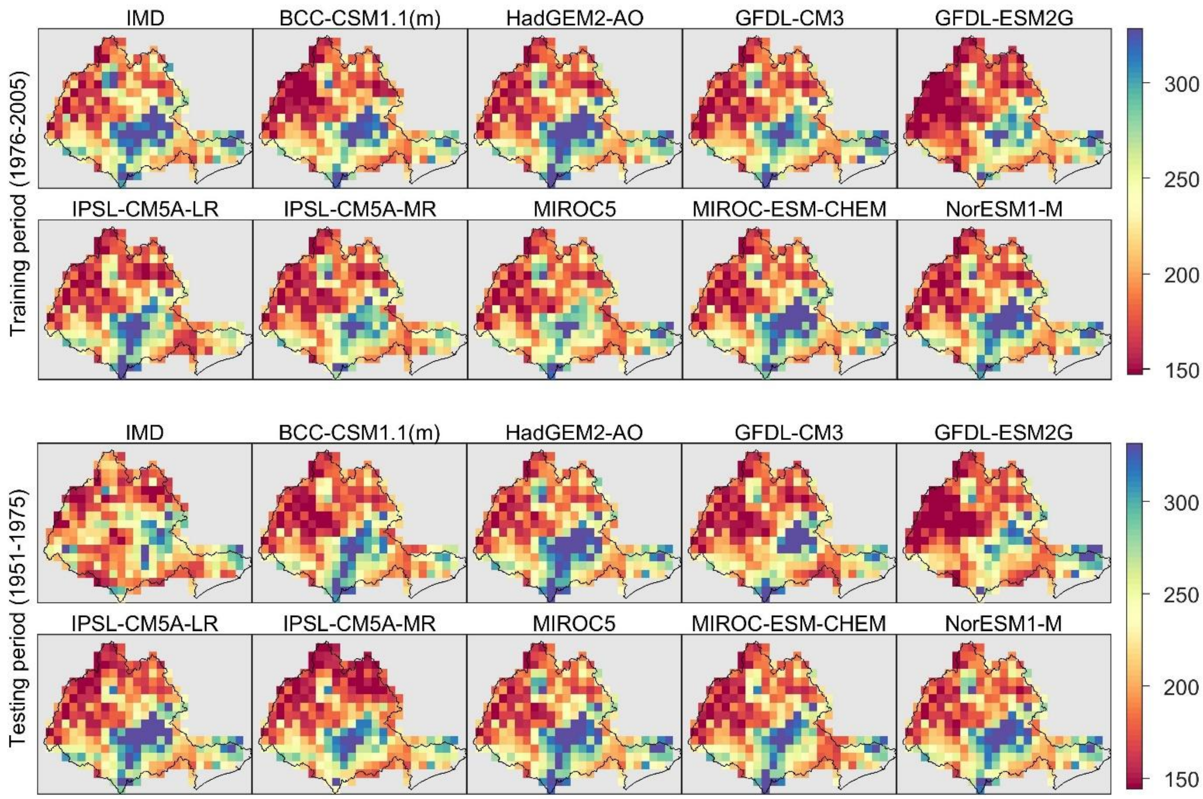


**Fig. 4** Spatial maps of 50 year return period rainfall for IMD and bias-corrected GCMs using monthly hybrid approach. Color bar in the right shows rainfall depth in mm.

Thus, it can be stated from the study that monthly hybrid approach can be utilized for quantile mapping over other two discussed techniques for improved extreme correction while preserving the seasonal climatology of rainfall.

## 4 Conclusions

The present study, proposes a novel approach of bias-correction i.e., monthly hybrid approach and compares it against two other approaches namely, seasonal approach and hybrid approach (proposed by Um et al., 2016) using a total of nine GCMs. Seasonal approach does not correct the extreme rainfall explicitly, therefore, does not perfrom better in capturing exremes during bias corection. On the other hand hybrid approach does not consider the seasonaility in rainfall time series, which is very critical in monsson dominated regions, hence, does not preserve the regional rainfall climatology or mean monthly distribution of rainfall adequately during bias correction.

The monthly hybrid approach which corrects the extreme rainfall using GEV distribution and applies quantile mapping using gamma distribution on monthwise daily time series was found to capture extremes with improved accuracy while capturing the rainfall seasonaility in close agreement with observation. Therefore, monthly hybrid approach should preferrably be used over other conventional bias-correction techniques while using GCMs rainfall projection for hydrological modeling or climate change impact analysis.

## Reference

Akhter, J., Das, L., Deb, A., 2017. CMIP5 ensemble-based spatial rainfall projection over homogeneous zones of India. Clim. Dyn. 49, 1885–1916. https://doi.org/10.1007/s00382-016-3409-8

Bisht, D.S., Sridhar, V., Mishra, A., Chatterjee, C., Raghuwanshi, N.S., 2019. Drought characterization over India under projected climate scenario. Int. J. Climatol. 39, 1889–1911. https://doi.org/10.1002/joc.5922

Dhage, P.M., Raghuwanshi, N.S., Singh, R., Mishra, A., 2016. Development of daily temperature scenarios and their impact on paddy crop evapotranspiration in Kangsabati command area. Theor. Appl. Climatol. https://doi.org/10.1007/s00704-016-1743-8

Hosking, J.R.M., 1990. L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics. J. R. Stat. Soc. Ser. B. https://doi.org/10.2307/2345653

Hosking, J.R.M., Wallis, J.R., 1997. Regional frequency analysis-An approach based on L-moments. Cambridge University Press, New York.

IPCC, 2007. Climate Change 2007: Synthesis Report, Intergovernmental Panel on Climate Change, Intergovernmental Panel on Climate Change.

Jacob, X.K., Bisht, D.S., Chatterjee, C., Raghuwanshi, N.S., 2019. Hydrodynamic Modeling for Flood Hazard Assessment in a Data Scarce Region : a Case Study of Bharathapuzha River Basin. Environ. Model. Assess. 1–18. https://doi.org/10.1007/s10666-019-09664-y Hydrodynamic

Jena, P.P., Chatterjee, C., Pradhan, G., Mishra, A., 2014. Are recent frequent high floods in Mahanadi basin in eastern India due to increase in extreme rainfalls? J. Hydrol. 517, 847–862. https://doi.org/10.1016/j.jhydrol.2014.06.021

Kumar, R., Chatterjee, C., 2011. Development of Regional Flood Frequency Relationships for Gauged and Ungauged Catchments Using L-Moments, in: In Extremis. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 105–127. https://doi.org/10.1007/978-3-642-14863-7

Kumar, R., Chatterjee, C., 2005. Regional Flood Frequency Analysis Using L-Moments for North Brahmaputra Region of India. J. Hydrol. Eng. 10, 1–7. https://doi.org/10.1061/(ASCE)1084-0699(2005)10:1(1)

Kumar, R., Chatterjee, C., Kumar, S., 2003. Regional flood formulas using L-moments for small watersheds of Sone subzone of India. Appl. Eng. Agric. 19, 47.

https://doi.org/10.13031/2013.12736

Li, H., Sheffield, J., Wood, E.F., 2010. Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching. J. Geophys. Res. 115, 1–20. https://doi.org/10.1029/2009JD012882

Pai, D.S., Sridhar, L., Rajeevan, M., Sreejith, O.P., Satbhai, N.S., Mukhopadyay, B., 2014. Development of a new high spatial resolution (0.25° × 0.25°) Long Period ( 1901-2010 ) daily gridded rainfall data set over India and its comparison with existing data sets over the region. Mausam 65, 1–18.

Reiter, P., Gutjahr, O., Schefczyk, L., Heinemann, G., Casper, M., 2017. Does applying quantile mapping to subsamples improve the bias correction of daily precipitation? Int. J. Climatol. https://doi.org/10.1002/joc.5283

Salvi, K., S., K., Ghosh, S., 2013. High-resolution multisite daily rainfall projections in India with statistical downscaling for climate change impacts assessment. J. Geophys. Res. Atmos. 118, 3557–3578. https://doi.org/10.1002/jgrd.50280

Samantaray, D., Chatterjee, C., Singh, R., Gupta, P.K., Panigrahy, S., 2015. Flood risk modeling for optimal rice planning for delta region of Mahanadi river basin in India. Nat. Hazards 76, 347–372. https://doi.org/10.1007/s11069-014-1493-9

Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res. Atmos. 106, 7183–7192. https://doi.org/10.1029/2000JD900719

Teutschbein, C., Seibert, J., 2012. Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. J. Hydrol. 456–457, 12–29. https://doi.org/10.1016/j.jhydrol.2012.05.052

Um, M.-J., Kim, H., Heo, J.-H., 2016. Hybrid approach in statistical bias correction of projected precipitation for the frequency analysis of extreme events. Adv. Water Resour. 94, 278–290. https://doi.org/10.1016/J.ADVWATRES.2016.05.021

Wilby, R.L., Wigley, T.M.L., 1997. Downscaling general circulation model output: a review of methods and limitations. Prog. Phys. Geogr. Earth Environ. 21, 530–548. https://doi.org/10.1177/030913339702100403